



# Avatar and Sense of Embodiment: Studying the Relative Preference Between Appearance, Control and Point of View

Rebecca Fribourg, Ferran Argelaguet Sanz, Anatole Lécuyer, Ludovic Hoyet

## ► To cite this version:

Rebecca Fribourg, Ferran Argelaguet Sanz, Anatole Lécuyer, Ludovic Hoyet. Avatar and Sense of Embodiment: Studying the Relative Preference Between Appearance, Control and Point of View. IEEE Transactions on Visualization and Computer Graphics, 2020, 26 (5), pp.2062-2072. 10.1109/TVCG.2020.2973077 . hal-02868067

**HAL Id: hal-02868067**

**<https://inria.hal.science/hal-02868067>**

Submitted on 22 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Avatar and Sense of Embodiment: Studying the Relative Preference Between Appearance, Control and Point of View

Rebecca Fribourg, Ferran Argelaguet, Anatole Lécuyer, and Ludovic Hoyet



Fig. 1. The four tasks implemented in the subjective matching experiment with the avatar's appearance at maximum level of realism. From left to right: *Punching*, *Soccer*, *Fitness* and *Walking*.

**Abstract**—In Virtual Reality, a number of studies have been conducted to assess the influence of avatar appearance, avatar control and user point of view on the Sense of Embodiment (SoE) towards a virtual avatar. However, such studies tend to explore each factor in isolation. This paper aims to better understand the inter-relations among these three factors by conducting a subjective matching experiment. In the presented experiment ( $n=40$ ), participants had to match a given “optimal” SoE avatar configuration (realistic avatar, full-body motion capture, first-person point of view), starting by a “minimal” SoE configuration (minimal avatar, no control, third-person point of view), by iteratively increasing the level of each factor. The choices of the participants provide insights about their preferences and perception over the three factors considered. Moreover, the subjective matching procedure was conducted in the context of four different interaction tasks with the goal of covering a wide range of actions an avatar can do in a VE. The paper also describes a baseline experiment ( $n=20$ ) which was used to define the number and order of the different levels for each factor, prior to the subjective matching experiment (e.g. different degrees of realism ranging from abstract to personalised avatars for the visual appearance). The results of the subjective matching experiment show that point of view and control levels were consistently increased by users before appearance levels when it comes to enhancing the SoE. Second, several configurations were identified with equivalent SoE as the one felt in the optimal configuration, but vary between the tasks. Taken together, our results provide valuable insights about which factors to prioritize in order to enhance the SoE towards an avatar in different tasks, and about configurations which lead to fulfilling SoE in VE.

**Index Terms**—Avatar, Sense of Embodiment, Immersive Virtual Reality, Psychophysics, Subjective Matching Technique

## 1 INTRODUCTION

The use of virtual avatars has become a striking feature in the latest developments of Virtual Reality (VR) applications. This increasing importance given to virtual avatars reinvigorates the research interests in the approaches to design them in such a way that users feel embodied. However, the design and conception of avatars is tailored by a number of technical (e.g. motion capture capabilities), data (e.g. 3D model reconstruction) and algorithmic (e.g. animation) constraints. Indeed, a fully functional avatar requires a vast amount of choices, and yet little is known about how the combination of choices are accepted by users, and affect their perception of the resulting avatars.

In the past years, many studies have tried to better understand how users perceive their avatar in VR by evaluating their Sense of Embodiment (SoE). More precisely, they focused on three subcomponents of the SoE [19]: the Sense of Self-Location, the Sense of Ownership and the Sense of Agency. From those researches emerged different “factors of influence” towards these three subcomponents, e.g., the avatar’s appearance [1] or the user’s point of view [16]. However, despite the worthwhile highlights brought by these studies, the inter-relations between the factors influencing the SoE remain uncertain. Indeed, if we start to better understand the influence of isolated factors on the SoE, we still have little information regarding the relative contribution of each factor towards the SoE, or regarding the user’s preference for a factor over another while being embodied in an avatar. As for today,

several questions remain open: Is there a dominant contribution between the factors of influence towards the SoE? Should some of these factors be prioritized in the creation of virtual avatars?

In order to provide insights to these questions, we present two experiments exploring user preference and perception of three factors commonly found in the literature to influence the sense of embodiment, namely the avatar’s visual appearance, the avatar’s control, and the user point of view. The first experiment (baseline experiment,  $n=20$ ) had the objective to create an ordered list for the levels within each factor (e.g., ranking between the different degrees of realism for an avatar appearance, ranging from abstract to personalised avatars). For each factor, participants experienced all levels while performing a task and had to rank the preference for each level in a scale from 0 to 100. The task consisted in recreating a yoga posture in front of a mirror.

The second experiment ( $n=40$ ) used the results obtained in the baseline experiment in order to explore through a subjective matching technique how participants combined them to reach a given level of SoE. Subjective matching experiments have already been successfully conducted on the factors impacting Place Illusion and Plausibility Illusion in VEs [41]. Such experiments aim at studying qualia, i.e. a quality or property as perceived or experienced by a person such as the Place Illusion, the Plausibility illusion or what interests us in this paper, the Sense of Embodiment, avoiding the use of subjective questionnaires or purely physiological and behavioral measures. More precisely, in our case the experiment consisted in having participants experiencing an “optimal” configuration of an avatar and then “recreate” the experienced SoE by iteratively increasing, one level at a time, one factor, starting from a “minimal” configuration. The final matched configuration, named accepted configuration, should match the same SoE experienced

• Rebecca Fribourg, Ludovic Hoyet, Anatole Lécuyer and Ferran Argelaguet: Inria, Univ Rennes, CNRS, IRISA, France.

with the “optimal” configuration. The initial “optimal” configuration was supposed to elicit a high SoE as it considered a partially customized avatar, full-body motion capture and a first-person point of view, while the “minimal” configuration consisted in a minimal avatar, with automatic animations and a third-person point of view. These configurations were defined according to ranking results from the baseline experiment. The choices of the participants provide insights about their preferences and perception over the three factors. In addition, to assess the potential impact of users actions while being embodied in an avatar, the subjective matching experiment considered four different tasks which covered four actions that can be done in a virtual environment: a) an interaction with the upper-body, b) an interaction with the lower-body, c) mimicking the actions of another virtual character full-body motions, or d) a constrained walking task. We had three main hypotheses. First, that we could create a monotonic ranking for the different levels of each factor. Second, that some factors would be prioritized over other factors. Finally, we expected the task to have an impact on the results.

Overall, our results validate our main hypotheses. First, a monotonic ranking for selected levels of each factor was successfully created. Second, it was shown that point of view and control levels were consistently increased by users before appearance levels. Third, several configurations were identified with equivalent SoE as the one felt in the optimal configuration, but tend to vary between the tasks. Taken together, our results give insights of which factors to prioritize to enhance the SoE towards an avatar, and about configurations which lead to SoE judged to be similar to the SoE experienced in the optimal configuration.

In summary, the main contributions of this paper are threefold. First, an experiment measuring the relative user preference of three factors related to the SoE: the avatar appearance, the avatar control and the user’s point of view. Second, the first subjective matching experiment assessing the relative contribution of these three factors on the sense of embodiment when performing four distinct tasks. Third, the highlight of valuable insights about which factors to prioritize in order to enhance the SoE towards an avatar in different tasks, and about configurations which lead to fulfilling SoE in VE.

## 2 BACKGROUND

In Virtual Reality, the Sense of Embodiment, as described by Kilteni et al. [19], refers to the feeling of being inside, controlling and having a virtual body, and can be decomposed in three respective and distinct subcomponents: the Sense of Self-Location, the Sense of Agency and the Sense of Ownership. In a similar direction, De Vignemont [8] considered that the SoE is the synthesis of three related subcomponents (spatial, motor and affective) and that it could be expressed on a continuous scale from no to full degree. Thus, theoretically, by altering the perception of each individual subcomponent it would be possible to alter the overall SoE. A notable number of researches therefore have explored the factors that could influence the different subcomponents. We focus on the three factors that have been more widely studied: avatar appearance, avatar control and the user’s point of view. In the first section, we focus on the main subcomponent that is influenced by each factor and in a second section we discuss potential inter-relations between the factors and other subcomponents.

### 2.1 Factors Influencing the Sense of Embodiment

The **appearance** of the avatar has been demonstrated to be a critical factor of influence to elicit the Sense of Ownership [1, 26]. The appearance of the avatar can be divided into several characteristics of influence: the general structure of the virtual body, the shape and dimension of body parts and the render style. These characteristics combined together contribute to different levels of avatar realism, anthropomorphism and fidelity towards the user’s real body, and by extent to different degrees of Sense of Ownership. For instance, Lin and Jörg [26] showed this sense was stronger with a more realistic human hand model compared to a non-anthropomorphic hand model. Similarly, while it is still possible to feel ownership towards full-body avatars with different degrees of anthropomorphism [28], the Sense of Ownership tends to be higher when the avatar clothes and skin tone match the user’s ones [30]. In a higher level of customization, the use of 3D scanned replicas has been

also considered [17, 48], and results have shown that they positively influence the sense of ownership. However, such approaches require complex 3D capture setups. Furthermore, if the latest researches mainly focused on exploring avatars with high realism and fidelity, other researches also explored the capability of users to feel ownership towards an avatar which differs from their self-representation in terms of body structure [23] or gender [42], showing that it is in general possible to elicit a Sense of Ownership towards such avatars.

The **control** of an avatar seems also to have a direct impact on user Sense of Agency. The actions performed by the avatar are judged by users, i.e. “did the avatar performed the action I wanted?”, referred as the judgment of agency, but also, in a pre-motor phase, “can the avatar perform the action I want?”, referred as the feeling of agency [44]. Indeed visuomotor congruence between real and virtual body movements highly contribute to the Sense of Agency [7, 41], while discrepancies between visual and motor information tend to decrease it [10, 37]. Regarding the feeling of agency, other studies showed that it is possible for users to feel illusory Sense of Agency towards actions they did not cause when some requirements are respected, such as a close match between users intentions and subsequent actions [21, 29]. Nagamine et al. also support the important role of motor control in the recognition of one’s own actions [32]. Regarding avatar animation techniques, such as the use of inverse kinematics or motion capture, some studies explored the influence of motion artifacts (latency, noise) in such techniques on the Sense of Agency, showing for instance that it impacts the Sense of Agency but does not break it [49]. Others also explored the impact of such controls on the Sense of Ownership [36] or on the Sense of Embodiment [33]. However, no studies explored to our knowledge the influence of the actual animation technique on the Sense of Agency.

The **point of view** (PoV) of users in the virtual environment with respect to their avatar determines the spatial relationship between their avatar and their virtual body. Such relationship can have an impact on where one perceives oneself to be located and thus alters the Sense of Self-Location. For example, a first-person PoV can easily induce the Sense of Self-Location [16], while a third-person PoV is more likely to reduce it [12, 31]. However, in out-of-body experiments, the illusion of self-Location might persist if it is preceded by a stimulation period [6]. More precisely, the presence of congruent visuotactile stimulation was also shown to be a key factor regarding the Sense of Self-Location, as it may lead users to mislocalize themselves towards the virtual body, to a position outside their bodily borders [25].

While some factors appear to be clearly linked to a specific subcomponent, their influence seems more complex, and some interrelations may exist within the factors and the subcomponents they influence. For this reason it is also interesting to have knowledge for each factor of its influence on all the subcomponents of the SoE.

### 2.2 Inter-relation Between the Factors Influencing the SoE

A number of studies have shown that one isolated factor can impact more than one subcomponent. For example, while the **appearance** factor seems to be mainly connected to the Sense of Ownership, it could increase the control expectations over the virtual body. For example, in the work of Argelaguet et al. [1] a virtual hand with lower realism elicited a stronger Sense of Agency over a realistic virtual hand. Authors hypothesized that the decrease in the Sense of Agency was due to the mismatch between the control mechanism and the actual appearance of the virtual hand. Users’ expectancy about the actual interactions capabilities of a realistic virtual hand were not met decreasing their Sense of Agency. On the same basis, while the **control** factor seems directly linked to the Sense of Agency, Steptoe et al. showed that the Sense of Ownership towards an extra body part (virtual human tail) was higher when users could actually control it by moving their hips [43]. Thus, suggesting that the mere fact of being able to control your virtual body has an effect on the Sense of Ownership. Similarly, the work of González Franco et al. [15] showed that being able to control the upper-body of the avatar elicited a higher Sense of Ownership that when just an animation was played. It was also shown by Kokkinara et al. [22] that multisensory congruence such as visuo-motor-tactile congruence enhances the Sense of Ownership

and that it can preserve the same Sense of Ownership between third-person and first-person PoV towards an avatar [13], even though in most cases the Sense of Ownership is higher in first-person PoV [16]. This highlights that the **point of view** factor is not only related to the Sense of Self-Location but that it can also influence the Sense of Ownership.

However, due to all possible inter-connections between the factors influencing the sense of embodiment and its subcomponents, it remains challenging to quantify their impact on the perceived Sense of Embodiment as a whole. An analogous question was raised by Kiltner et al. [19] regarding the relationship between SoE and its subcomponents. Some research for instance would place the Self-Location as the most important subcomponent [5] while others would suggest the Sense of Ownership to be of low significance [34] and the Sense of Agency to be of much importance [45]. Overall, Kiltner et al. [19] insisted on the lack of current knowledge regarding the weight of each subcomponent contribution to the SoE, which coincides with the gap of knowledge regarding the importance of each factor regarding their influence on the SoE as an entire complex entity and not towards its specific subcomponents.

Common studies exploring the influence of factors towards the SoE usually focus on one factor at a time and measure its influence on the SoE with the use of subjective questionnaire [14] or purely physiological and behavioral measures [2, 9]. However, such measures do not allow the assessment of inter-relations between the factors influencing the SoE. Indeed, the assessment of this kind of inter-relations is challenging in terms of experimental protocol due to the numerous amount of possible factor combinations. To this respect, our paper aims to better understand the inter-relations among these three factors by conducting a subjective matching experiment which has already been used in the context of Presence [3, 4, 40, 41] and enables to evaluation of a relatively high number of factor combinations.

### 3 OVERVIEW AND GENERAL EXPERIMENTAL DETAILS

The main objective of this paper was to identify potential preferences within factors of influence towards the SoE. To do so, we first conducted a baseline experiment to define the number and order of the different levels for each factor of influence towards the SoE. We then conducted a subjective matching experiment, similarly to the studies on Presence of Slater et al. [41] and Skarbez et al. [40], in order to better understand the inter-relations between these factors. In this section, we detail the subjective matching technique used in our main experiment as well as the experimental details common to both experiments.

#### 3.1 Subjective Matching Technique

The subjective matching technique is a method commonly used in color science where a particular color sensation is considered as an equivalence class over a number of different wavelength distributions. Typically, users are presented with a color, then asked to reproduce the same color by additively mixing the three primary colors.

In the context of this paper, a particular SoE could similarly be considered as an equivalence class over different levels of factors that may influence it, and users were therefore asked to reproduce a given SoE by combining different levels of these factors. A combination of several levels of factors is called hereafter “configuration”. In our case, these factors are the Appearance, Control and Point of View, leading to numerous possible avatar configurations with many potential degrees of SoE. Moreover, the SoE felt in a specific configuration combining the three factors might be equivalent to one felt in another configuration of these factors. The subjective matching technique used in the experiment therefore involves users trying a specific “optimal” configuration of avatar, and remembering their SoE in this configuration. They are then asked afterwards to combine several levels of factors to match again the SoE felt in the initial configuration. More precisely, to each factor is associated a number of levels of improvement, assuming that having all the factors at their maximum level would lead to the best configuration in which users are more likely to have the highest SoE. This method therefore enables to highlight a) which factors participants are more likely to improve and in which order, and b) which configurations will elicit a SoE equivalent to the one felt in the best configuration.

#### 3.2 Factors and Levels

To do such an experiment, we chose to focus on three factors (independent variables), with the objective of covering as much as possible the different degrees of SoE likely to be felt towards an avatar. The visual *Appearance* of the avatar was chosen to encompass visual feedback of the avatar that relates to graphical features. The *Control* was chosen to embrace any capabilities of having the avatar animated in the VE. Finally, the *Point of View* was chosen to include different perspectives taken from a user towards the virtual body of the avatar. For each factor several levels were identified with an initial pre-supposed ranking which was refined in a baseline experiment (see Section 4). The main requirements for choosing the factors and levels were to ensure good coverage of potential implementations of an avatar according to each factor, as well as allowing the combination of levels between factors. For instance, we did not separate *Appearance* into texture and shape as realistic textures would hardly be combinable with abstract geometrical representations. Similarly, we did not include finger animation since it could not consistently be combined with all the appearance levels. These implications are discussed in more details in Section 7.

##### 3.2.1 Appearance

The appearance of an avatar can be addressed over several characteristics: the general structure of the virtual body, the shape and dimension of body parts, the render style, etc. Those characteristics combined together contribute to different levels of avatar realism, anthropomorphism and fidelity towards the user. As detailed in Section 2, many visual configurations of avatars have been tested in order to evaluate their influence on the SoE and more precisely on its subcomponents. For our experiment, we have selected 6 levels that we believed were the most represented in past studies (see Figure 2), ranging from low to high realism and anthropomorphism representations (including the distinction of three realistic avatars in terms of fidelity):

- ( $A_a$ ) Abstract avatar. Only extremities of the body are visually represented with white spheres.
- ( $A_b$ ) Stickman. Extremities and main body joints are visually represented with white spheres and cylinders.
- ( $A_c$ ) Dummy avatar. An avatar with a human body shape but a robotic appearance.
- ( $A_d$ ) Opposite realistic avatar. A realistic gender-matched humanoid avatar that participants chose among a list of 20 different avatars (20M, 20F) (see supplementary material) with the instruction of choosing one that they considered to be their opposite in terms of resemblance.
- ( $A_e$ ) Neutral realistic avatar. A realistic gender-matched humanoid avatar that participants chose among a list of 20 different avatars (20M, 20F) with the instruction of choosing one that did not evoke them anything particular.
- ( $A_f$ ) Personalized realistic avatar. A realistic gender-matched humanoid avatar that participants chose among a list of 20 different avatars (20M, 20F) with the instruction of choosing one that they considered to resemble them the most. This avatar could then be slightly personalized in terms of hair, eye and clothes color.

##### 3.2.2 Control

Similarly, we selected four levels of Control based on previous works, that we believed were most likely to have different effects on the SoE.

- ( $C_a$ ) Automatic animation. When participants enter a specific zone in order to perform the task, an animation is automatically launched on the virtual body which makes the avatar do the task while the participants actually have no control over it.
- ( $C_b$ ) Triggered animation. Pressing a button, participants can trigger themselves the animation performing a task in the VE (same animation as in  $C_a$ ).
- ( $C_c$ ) Inverse Kinematics. The virtual body is animated using Inverse Kinematics, enabling the animation of the avatar from participants’ head, hands and feet positions and orientations.
- ( $C_d$ ) Motion capture. The virtual body of the avatar is animated using a motion capture system (Xsens system).

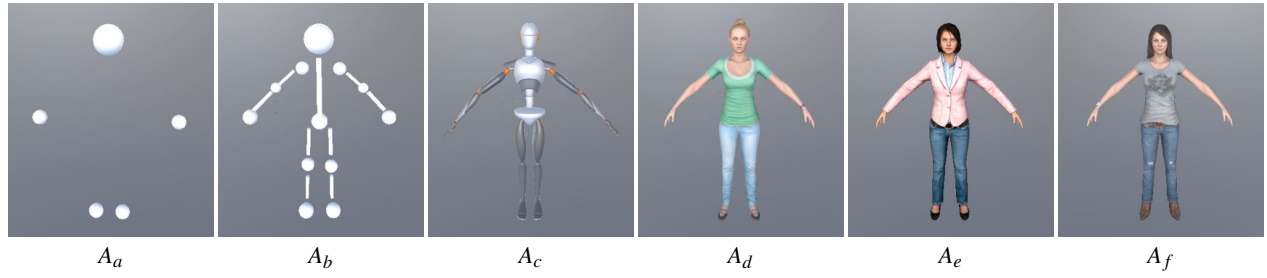


Fig. 2. Levels of the *Appearance* factor. From left to right: ( $A_a$ ) Abstract avatar, ( $A_b$ ) Stickman, ( $A_c$ ) Dummy avatar, ( $A_d$ ) Opposite realistic avatar, ( $A_e$ ) Neutral realistic avatar and ( $A_f$ ) Personalized realistic avatar.

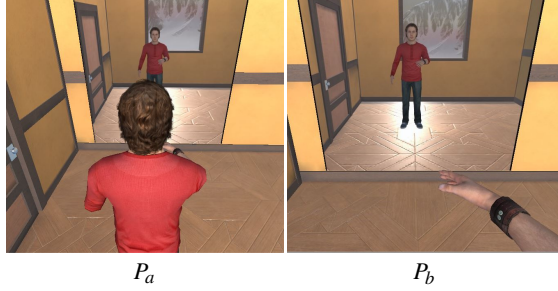


Fig. 3. The two levels of the *Point of View* factor: ( $P_a$ ) Third-person PoV, ( $P_b$ ) First-person PoV

### 3.2.3 Point of View

Two levels were chosen for the PoV depending on participants perspective towards the virtual body (see Figure 3).

- ( $P_a$ ) Third-person PoV. Users see their virtual body from a classical over-the-shoulder PoV, as commonly used in video games.
- ( $P_b$ ) First-person PoV. Users see their virtual body as if they were in the avatar's head (as they would see their own body in real life).

### 3.3 Apparatus

For both experiments, the virtual environment was developed in Unity (version 2018.3.14f1) and displayed using an HTC Vive PRO Head-Mounted-Display (HMD). For head tracking, the internal tracking of the HTC Vive HMD was used. For body tracking, participants wore an IMU-based (Inertial Measurement Unit) motion capture system (Xsens). IMU sensors were equipped on the participants using motion capture suit and straps. The body tracking was handled by the Xsens MVN Animate software platform and streamed to Unity in real time. When using Inverse Kinematics, the FinalIK plugin was used to animate the avatar by following the feet, hand and pelvis positions provided by the Xsens software. Participants also hold Vive Controllers in their hands to interact with the virtual environment.

### 3.4 Participants

Twenty participants took part in the baseline experiment (17 males and 3 females; mean/S.D. age:  $25.8 \pm 5.6$ ). Forty participants (20 males, 20 females; mean/S.D. age:  $32.5 \pm 10.1$ ) were recruited for the subjective matching experiment. For both experiments, participants were recruited from the university campus, were naive with respect to the purpose of the experiment and had normal or corrected-to-normal vision. The studies conformed to the declaration of Helsinki.

Before each experiment, participants were first briefed about the experiment, signed an informed-consent form and completed a demographic questionnaire. After this process, they were equipped with the Xsens motion capture system before undergoing a calibration procedure that would ensure the efficiency of the motion capture system but also allow to resize the avatar to participants dimensions. Finally they were equipped with the HTC Vive PRO HMD and started the experiment.

## 4 BASELINE EXPERIMENT

While previous work findings enabled us to pre-select and pre-rank several levels for each factor, little is actually known about the relative differences between all these levels in terms of their influence on the SoE. For instance, does a stickman actually elicit a significantly lower SoE compared to a dummy character? Or do animations driven by Inverse Kinematics elicit similar or lower levels of embodiment than animations driven by a motion capture system? To better measure these differences, and therefore provide significant levels of improvements between levels in the following subjective matching experiment, we decided to conduct a baseline experiment to accurately define the number and order of the different levels based on user preferences.

### 4.1 Experimental Protocol

The experiment consisted in making participants try and rate all the levels of each factor on a score between 0 and 100. To that aim, participants were immersed in a virtual environment representing a fitness room, facing a mirror and had to perform a task while testing all the levels of each factor. The task consisted in performing an easy yoga pose in the context of a “virtual yoga class” (see Figure 4, left). More precisely, a specific zone in front of the mirror was highlighted by a luminous disc on the ground, and the task consisted in going to this zone, doing the yoga pose, and going back to the initial position. The experiment was divided in three blocks, each corresponding to a particular factor, presented in random order. When a given factor was being tested, the two other factors were set at their pre-supposed maximum level (i.e  $A_f$ ,  $C_d$  or  $P_b$ ). In each block, participants started with the factor tested at a random initial level. A virtual slider, as well as virtual cubes next to the slider corresponding to each level of the factor (see Figure 4, center and right), were visible on their left. The order in which the virtual cubes were initially presented was also randomized and the cube corresponding to the random initial level tested was highlighted as being selected. Participants were instructed to proceed as follow. First, perform the task. Second, rate the level by positioning the virtual cube on the slider according to their preference in order to perform the task, ranking it simultaneously by its position relative to the other levels. Third, select another virtual cube with their controller in order to change the level of the factor. They had the possibility to try one level several times when needed to adjust their ratings. When all the cubes were positioned on the slider, the next block could start. The baseline experiment, including the welcoming of participants and consent form signing lasted about thirty minutes.

### 4.2 Recorded Data

There were two dependent variables in the baseline experiment for each factor: the score attributed to each level (from 0 to 100) and the ranking of these same levels between each other (from 0 to 5 for *Appearance*, 0 to 3 for *Control* and 0 to 1 for *PoV*).

### 4.3 Results

For the scores analysis, both the normality and homogeneity of variances assumptions were verified for *Appearance* and *Control*, with respectively the Shapiro-Wilk's Normality test ( $p = 0.3009$  for *Appearance*,  $p = 0.9766$  for *Control*) and Bartlett test ( $p = 0.3994$  for *Appearance*,  $p = 0.1569$  for *Control*). For *PoV* the homogeneity was verified



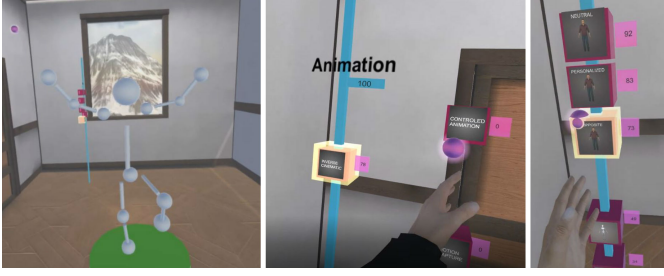


Fig. 4. Baseline experiment. Left: user performing the yoga pose with the Stickman appearance. Center and Right: user interacting with the scale to rate select and rate a given level of Control factor.

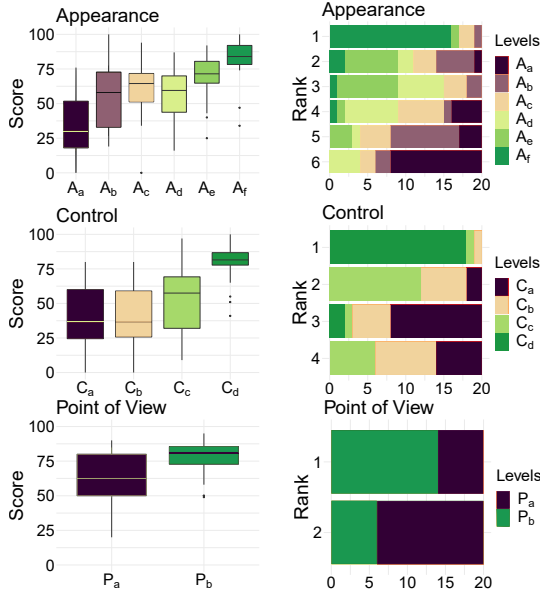


Fig. 5. Factors mean scores per levels (left) and rank distribution among participants (right).

( $p = 0.1159$ ) but not the normality. A one-way ANOVA analysis was thus performed for *Appearance* and *Control* and showed significant differences between mean scores of levels ( $[F_{1,95}=425.72, p < .0001]$  for *Appearance*,  $[F_{1,57}=232.57, p < .0001]$  for *Control*). Tukey's post-hoc tests ( $\alpha = .05$ ) were conducted to check significance for pairwise comparisons. For *Appearance*, A<sub>a</sub> was scored significantly lower than all other levels ( $p < .05$ ). A<sub>b</sub>, A<sub>c</sub> and A<sub>d</sub> were all only significantly lower than A<sub>f</sub> ( $p < .05$ ). For *Control*, C<sub>a</sub>, C<sub>b</sub> and C<sub>c</sub> were all significantly lower than C<sub>d</sub> ( $p < .0001$ ). Wilcoxon tests were conducted for *PoV* and showed that the P<sub>a</sub> was scored significantly lower than P<sub>b</sub> ( $p < .05$ ). The mean scores of levels for each factor are represented in Figure 5, left.

Regarding the ranking analysis, the normality and homogeneity of variances assumptions were not verified leading to an analysis for non parametric data. Wilcoxon tests showed that for *Appearance*, A<sub>a</sub> was ranked significantly below all the other levels ( $p < .05$ ). A<sub>b</sub> was ranked significantly below A<sub>e</sub> and A<sub>f</sub> ( $p < .05$ ). A<sub>c</sub> was ranked significantly below A<sub>f</sub> ( $p < .0001$ ). A<sub>d</sub> was ranked significantly below A<sub>e</sub> and A<sub>f</sub> ( $p < .05$ ). Finally, A<sub>e</sub> was ranked significantly below A<sub>f</sub>. Regarding *Control*, C<sub>a</sub> was ranked significantly below C<sub>c</sub> and C<sub>d</sub> ( $p < .05$ ). C<sub>b</sub> was ranked significantly below C<sub>d</sub> ( $p < .0001$ ). C<sub>c</sub> was ranked significantly below C<sub>d</sub> ( $p < .0001$ ). Regarding the Point of View, P<sub>a</sub> was ranked significantly lower than P<sub>b</sub> ( $p < .05$ ). The distributions of each level per rank for each factor are represented in Figure 5, right.

#### 4.4 Discussion and Levels Selection

The aim of this experiment was to better understand user preferences and relative ratings of the different levels for each of our factors. Pre-supposed orders had been hypothesized and were partially supported by the results of scores and ranking.

Table 1. Levels selected for the subjective matching experiment

Appearance	Control
(A = 0) Abstract avatar	(C = 0) Automatically launched animation
(A = 1) Stickman	(C = 1) Inverse Kinematics
(A = 2) Neutral realistic	(C = 2) Motion Capture
(A = 3) Personalized realistic	
Point of View	
(P = 0) Third Point of View	
(P = 1) First Point of View	

For *Appearance*, the Abstract avatar was highlighted as the lowest level and the Personalized avatar as the highest level among all. The Neutral avatar was ranked significantly lower than the Personalized Avatar and higher than the Stickman and Dummy. However, the scores and ranking between the Opposite and Neutral avatars do not permit to clearly rank one above the others. Another block with close scores and rankings appeared between the Stickman, the Dummy, and the Opposite avatars making it hard to place one above the other. However, among this block, the Stickman was the only one ranked significantly below the Neutral avatar. While those results were mostly expected, it is surprising that the Dummy avatar was rated that close to the non personalized realistic avatars, since it is usually shown that the more realistic the avatar is, the higher the Sense of Ownership towards it is [1]. However, this result is in line with Lugin et al.'s study [28] showing similar levels of body ownership, as well as enjoyment, towards avatars with different levels of anthropomorphism (e.g. robot and realistic avatars). Moreover, while the higher score given to Personalized avatars compared to Opposite and Neutral ones is in line with Waltemate et al. study [48], interestingly no significant difference was found in the scores between the Opposite and Neutral avatars. It is nevertheless hard to interpret this result due to the variability of these two levels: participants were choosing themselves these avatars in a global list of 20 avatars per gender. For *Control*, the results only highlighted that Motion Capture was scored and rated higher than all the other levels, placing the Automatic animation, Triggered animation and Inverse Kinematics in the same block. However, among this block, only the Automatic animation was significantly ranked below Inverse Kinematics. While we did not expect such a difference in ratings between the Motion Capture and the Inverse Kinematics, we believe that the fact that the avatar was realistic while testing this factor may have allowed to see more easily potential motion artefacts. For *Point of View*, both scores and ranking results showed a preference for the first-person PoV against the third-person PoV, which is consistent with previous work [16].

In addition to gaining insights about user preferences, one of the goals of this baseline experiment was to define an ordered and rated subset of levels for each factor to be used in the subjective matching experiment. This seemed particularly important as the subjective matching experiment presented in the following section required participants to select the next factor to improve in order to increase their SoE, with the goal of matching a previously experienced level of SoE. As introducing levels which were not different enough within a factor might have introduced a bias towards selecting one factor in priority over another, the levels that were not significantly scored or ranked between each other were eliminated from the subset. Therefore, for *Appearance*, the Dummy and the Opposite avatar were removed giving a final ordered and rated subset of: Abstract, Stickman, Neutral and Personalized avatars. Proceeding similarly to the elimination of *Appearance* levels, the Triggered animation was thus removed from the *Control* subset giving the final ordered and rated subset: Automatic animation, Inverse Kinematics, Motion Capture. For the *Point of View*, both levels were thus kept in that order: Third-person PoV, First-person PoV. Table 1 summarizes the final levels selected for the subjective matching experiment.

#### 5 SUBJECTIVE MATCHING EXPERIMENT

The goal of this experiment was to study the relative contribution of the *Appearance*, *Control* and *Point of View* factors towards the SoE,

using the pre-selected levels for each factors obtained from the Baseline experiment. In other words, do users have preferences between those factors when it comes to enhance their SoE towards an avatar?

## 5.1 Tasks

Potential preferences regarding factors influencing the SoE may depend of the task performed in the VE. Indeed, the way users interact with the virtual environment may induce them to look more or less to certain parts of their virtual body, or more generally to pay more or less attention to their virtual body. The presence of collisions between the virtual body and the VE leading to visible feedback of changes in the VE may also influence the perception of the virtual body and thus the SoE. More abstractly, the general context of the interaction, its gamification [46, 50] or social aspect [38] might influence on users perception towards the overall VE.

For these reasons we hypothesized that the type of action performed by users in the VE would influence the SoE, and therefore designed four different tasks with the goal of covering a wide range of actions that an avatar can do in a VE. First, we designed two tasks involving direct interaction between the virtual body and the VE, one involving the upper-body and one involving the lower-body. Second, we designed a task involving no direct interaction between the virtual body and the VE, but the presence of another virtual character. Finally, we designed a walking task, navigation being a main and one of the most common interaction task in VR. We describe the tasks more in detailed hereafter:

- The *Punching* task consisted in hitting a punching bag, involving the virtual upper-body to be interacting directly with the VE (see Figure 1, first).
- The *Soccer* task consisted in kicking a soccer ball, involving the virtual lower-body to be interacting directly with the VE (see Figure 1, second).
- The *Fitness* task consisted in following fitness movements instructed by a “fitness teacher” (see Figure 1, third).
- The *Walking* task consisted in walking straight while avoiding obstacles on the floor. Low walls constrained the direction of the path to walk on (see Figure 1, fourth).

These four tasks were entered in the same general context of a fitness scenario, and participants were immersed in a virtual fitness room in front of a virtual mirror. Participants started on a circular green carpet, and always moved towards another green carpet in front of them to perform the task. The levels of each factor were also the same for the four tasks, with the unique difference that the actual animation of  $C_0$  (Automatically launched animation) was tailored for each task. For the *Punching* task, the automatic animation made the dominant hand punch the punching bag once, while for the *Soccer* task it made the dominant foot kick the ball. For the *Fitness* task the automatic animation displayed the same fitness movements shown by the virtual teacher. Finally, for the *Walking* task the automatic walking solution from FinalIK was applied to animate the feet so that they avoided obstacles when collisions were close, i.e., to step over the obstacles. The automatic animations specific to each task are presented in the accompanying video for illustrative purposes.

A mixed design was chosen for the experiment. Each participant performed randomly only two tasks. This choice was done to reduce experiment duration time and to ensure the engagement of the participants. The design ensured that each task was performed by 10 male and 10 female participants, the order of the tasks was counterbalanced.

## 5.2 Experimental Protocol

Participants started the experiment with a first exposure which had a threefold objective. First, it enabled participants to become familiar with the VE and the tasks to perform. Second, they were instructed to test and become familiar with all the possible levels of each factor. Finally, they then performed the tasks with the best avatar configuration (i.e., with the highest level for each factor:  $\{3, 2, 1\}^1$ ), and in that case

<sup>1</sup>Notation  $\{i, j, k\}$  represents an avatar configuration with levels  $A_i, C_j, P_k$

were instructed to focus on their SoE towards the avatar. Considering that the notion of “Sense of Embodiment” was not instinctive to understand for participants, we detailed the instruction to participants based on the description made in Kiltner et al. work [19]: “Please be aware of your SoE towards your virtual body while doing the task, considering your SoE as a union of the feeling of ownership you have towards the virtual body, the feeling of control you have over it, and the feeling of being spatially located in this virtual body”. After making sure that participants had tested all the improvements they could do towards the virtual avatar, and had memorized their SoE in the best configuration for the tasks, the second part of the subjective matching procedure started. Participants were instructed beforehand that for each task, they would perform several trials in which they would start in a low level configuration of avatar, with the goal of reaching the same SoE they had experienced in the “optimal” configuration. The initial configuration could either be all the factors at level 0 ( $\{0, 0, 0\}$ ) or just one factor at level 1 ( $\{0, 0, 1\}, \{0, 1, 0\}, \{1, 0, 0\}$ ). Each participant started once with each configuration giving 4 trials per task. In order to minimize ordering effects, the order of the starting configurations for each task was counterbalanced following a Latin square design.

Participants then increased a factor by telling the experimenter which factor they wanted to improve. Similarly, they were also instructed to notify the experimenter when their SoE matched the one felt in the “optimal” configuration of avatar. However, participants were asked to keep on making choices to improve the factors until they had reached the final configuration, even if the match happened before reaching the “optimal” configuration.

After completing all the trials for the two tasks, participants completed a post-experiment questionnaire, including the standardized embodiment questionnaire [14], the SUS presence questionnaire [47], as well as a series of questions to rate the factors regarding their preference when improving their avatar. While participants were asked to answer the presence questionnaire and rate the factors focusing on the general experiment (including both tasks), they were instructed to answer the embodiment questionnaire thinking of the avatar in the latest task tested, for which they had matched the high SoE. The whole experiment, including welcoming of participants, reading and signing the consent form, and answering questionnaires lasted around one hour.

## 5.3 Recorded Data

The recorded data includes participants choices during the experiment as well as the answers to the post-experiment questionnaire. First, there is the “Accepted Configurations”, i.e. the configurations at which participants declared to feel an equivalent SoE compared to what they felt in the “optimal configuration”. Second, there is the transitions set, meaning the order of improvements made by participants to go from one configuration to another. Finally, there are the answers to the embodiment and presence questionnaire (respectively 7-point and 5-point Likert scale) as well as the ratings made by participants regarding their general preference of factors (7-point Likert scale), all collected from the post-experiment questionnaire.

## 6 RESULTS

In this analysis we made the same assumption than Slater et al. [41] and Skarbez et al. [40], namely that the results for each repetition are statistically independent. Since there were performed by the same participant, they are not truly independent, but each trial started with a different initial configuration, forcing participants to reconsider their first choices each time. In this section, we report our analysis according to three measures: the identified Accepted Configurations, the transitions made by participants from the initial configuration to the optimal one, and finally their responses to the post-experiment questionnaire.

### 6.1 Accepted Configurations

To analyse the results concerning the Accepted Configurations, we first computed separately for each task the probability of accepting a configuration (Figure 6, top), which corresponds to the number of times participants reported a match of SoE for a given configuration

over the total number of accepted configurations (4 trials  $\times$  20 participants = 80 accepted configurations in total). If there was no match before the optimal configuration, this configuration was considered as the Accepted Configuration. For example, in the *Punching* task, the configuration {1,2,1} was accepted 9 times, which thus represents 11% of the total accepted configurations. We can observe that configuration {1,2,1} was the most accepted configuration for all tasks except *Punching*, for which the most accepted configurations are spread between configuration {2,2,1} and {3,2,1}.

Second, we computed for each task the conditional probability of participants reporting a match when experiencing a configuration (Figure 6). For instance, the configuration {1,2,1} in the *Punching* task was attained 35 times, while a match was only reported in 9 trials, meaning that there is a 26% probability for participants to report a match when attaining this configuration. Results are overall in line with the global probabilities computed, but also give additional information regarding configurations that may not have been often reached, but were mostly accepted when they were. For instance, in *Fitness* and *Walking*, configuration {3,1,1} was only reached 12 and 6 times, but when they were, they had more than 75% chance to be accepted.

Third, we computed for each task the probability of accepting a configuration depending on the participants' gender (see Figure 9), since several studies already showed that the perception of the virtual environment [40] and avatar [39] may vary accordingly. We can observe differences between males and females in *Punching* and *Walking*. In both tasks while males mostly accepted configurations {2,2,1} (44%) and {1,2,1} (45%) respectively in *Punching* and *Walking*, women tended to need higher level of appearance by accepting in majority configuration {3,2,1} (46% in *Punching* and 53% in *Walking*).

## 6.2 Transitions

A transition probability matrix was constructed with the configurations chosen by the participants. Since all participants were asked to improve the configurations until the optimal configuration, there were 6 improvements for each trial starting in configuration {0,0,0} and 5 improvements for the other trials. This makes a total of 21 improvements per participants per task, and a global total for all participants and all

tasks of 1680 improvements. This matrix enabled us to compute the probability distribution over the configurations for any given configuration, and the elaboration of a Markov chain for each of the four tasks (Figure 7). Each graph represents the probability distribution for each possible transition (configurations most explored are represented in green, while those barely explored are represented in red). The most likely path were also identified for each task and presented in Figure 8.

Over all tasks, results show that a clear majority of participants preferred to increase first their level of *Control* or *Point of View* against their *Appearance*. When the first choice was to improve either the *Control* or *Point of View*, the second decision was mostly to improve the other one next, leading to configuration {0,1,1}. At that point, in all tasks except *Soccer* most participants tended to improve their appearance ({1,1,1}), except for the *Soccer* task where the next choice was in majority to increase again the level of *Control* ({0,2,1}). Afterwards, participants mostly attained the same configuration {1,2,1}, by increasing the *Appearance* in *Soccer* or the *Control* in the other tasks. From this configuration, only the *Appearance* could be further increased until the final configuration {3,2,1}.

## 6.3 Post-experiment Questionnaire

From the Presence and Embodiment questionnaires we computed the mean scores for Presence regarding the global experiment ( $4.70 \pm 0.89$  (S.D.)) and Embodiment for each task (*Punching*:  $5.07 \pm 0.69$ , *Soccer*:  $5.23 \pm 0.80$ , *Fitness*:  $5.04 \pm 0.51$  and *Walking*:  $5.26 \pm 0.75$ ). Kruskal-Wallis tests were performed on embodiment scores showing no significant differences between tasks.

Moreover, mean scores of preference were computed for each factors (see Figure 10). Friedman tests showed significant differences between factors for the mean scores of preference attributed to each ( $p < .001$ ). Wilcoxon tests were thus conducted, showing that *Control* and *Point of View* were both rated on average significantly higher in terms of preference in order to improve the avatar ( $p < .001$ ).

## 7 DISCUSSION

### 7.1 Importance of Point of View and Control

According to our results, the Point of View and the Control clearly appeared as the preferred factors when improving the configuration of the avatar. This is primarily reflected in the first transition made by most of the participants in all tasks, as they chose to increase first either their level of Control or their Point of View at least 90% of the time whatever the task (90% of the time in *Punching*, and 95% for *Soccer*, *Fitness* and *Walking*). This is also visible in the most likely paths, where increases of the appearance level typically happen late in the paths. The preference regarding control and point of view over appearance is also notable in the most accepted configurations where nevertheless, some differences are to be noted among the tasks. Indeed, the configuration accepted most was {1,2,1} in *Soccer*, *Fitness* and *Walking*, while {2,2,1} and {3,2,1} were the most accepted in *Punching*. It thus seems that for all tasks except *Punching* a low level of appearance (stickman) was enough to match the level of SoE felt in the optimal configuration, while interestingly a higher level of appearance was required in *Punching*. In addition, in the post-experiment questionnaire *Control* and *Point of View* were rated significantly higher than *Appearance* (see Figure 10). Overall, these results underline a lower popularity of the appearance factor compared to the control and point of view. It is a rather intriguing result since the appearance of an avatar is a factor widely studied in VR and known to have a strong impact on the sense of body ownership. For instance, Lugin et al. [27] and Latoschik et al. [24] showed that more realistic avatars elicit higher sense of body ownership. Yet, when participants have to choose between the appearance and other factors in order to improve their avatar (with the goal of reaching an equivalent SoE as the one felt in the optimal configuration), they tend to depreciate the importance of the appearance in the improvement process. We may wonder in that case if the control and point of view influence the way the appearance of an avatar is perceived. While potential interrelations between the control and appearance of an avatar have partially been explored in the context of co-presence [18], this question remains open when concerning the SoE. In addition, these results also echo the

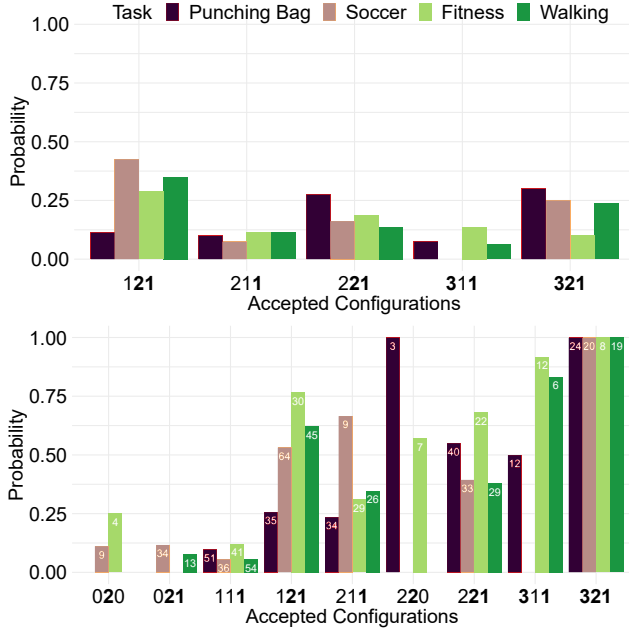


Fig. 6. Probability for configurations to be accepted (top) and conditional probability for configurations to be accepted if reached (bottom). The number written on each bar represents the number of times the configuration was reached. The levels of factors are in bold format when at their maximum. For readability purpose, only configurations with a probability of acceptance higher than 10% are shown.



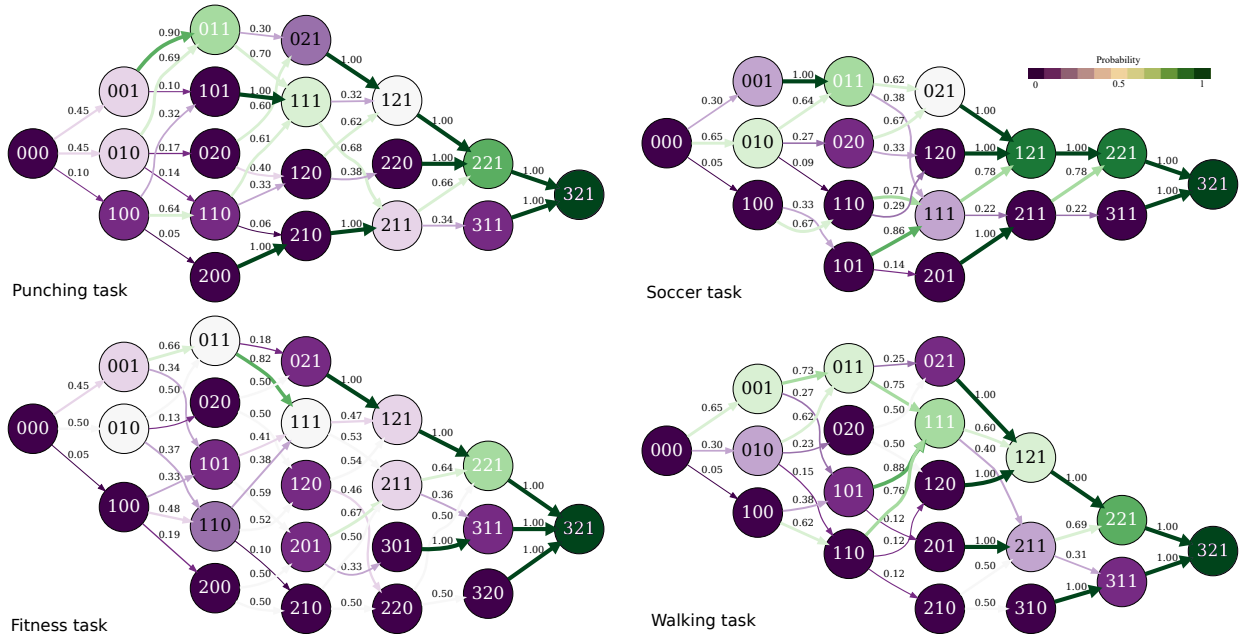


Fig. 7. Markov chains representing the transition matrix probability for each task. The color of a node represents the probability that the node is reached. The color and the thickness of the edges represent the transition probability from a given node.

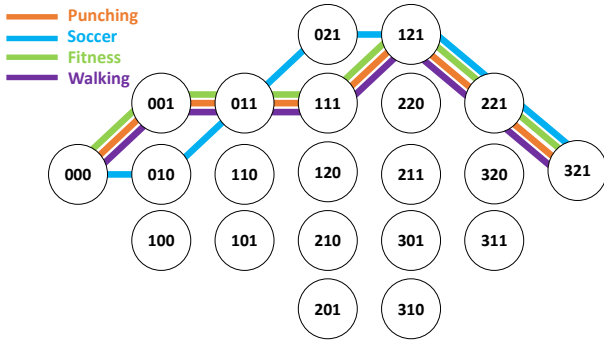


Fig. 8. Most likely path for all four tasks.

concern raised by Kiltani et al. [19] regarding the lack of knowledge about the contribution weight on the SoE of its subcomponents. While this question remains open, we hope our research will serve as a basis for further studies on the subject.

## 7.2 The Control Rush

While the preference attributed to control was reflected in all tasks, an increased interest was especially given to it in the soccer task. This is shown in participants' first choices of improvement: while control and point of view were equally increase first in *Punching* and *Fitness*, control was increased twice more often than point of view in *Soccer*. Furthermore, the most likely paths also highlight the preference of the control at a second stage. While in other tasks the most likely choice was to improve the appearance from configuration {0,1,1} (abstract avatar, IK and first-person PoV), control was mostly chosen instead in *Soccer*. This result is interesting since it shows that in this task, even with very low visual appearance (abstract avatar: only head and extremities represented), the control was increased at its maximum level (from IK to motion capture). This is a rather intriguing result since the major improvement made from IK to motion capture is the gain in precision regarding the position and orientation of middle parts of the body (knees, elbows, etc.), which is not visible with the abstract avatar. We may then wonder why the control was that much improved in that task since the main change between the two levels of control should not have been visible. While our results do not allow to answer this question, it is important to consider the potential influence of the

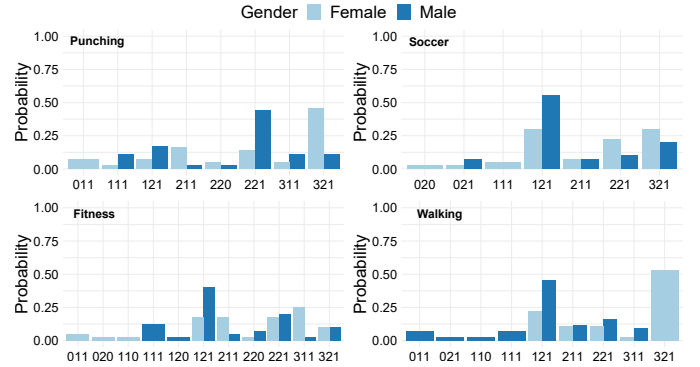


Fig. 9. Probability of a configuration to be accepted per task and depending on participants gender.

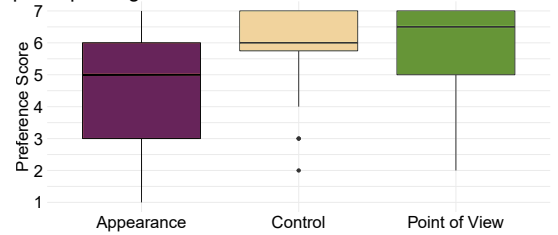


Fig. 10. Mean scores from the post-experiment questionnaire according to users' preference of improving the given factor on the avatar.

task characteristics. For instance, while participants were precisely instructed that the objective of the task was not to score a goal but only to kick the ball, whether the ball entered or not the goal could still have been interpreted as a success or failure by participants. A possible explanation could therefore be that participants associated the increased level of control with an increased chance of scoring a goal, inciting them to further increase this factor first.

## 7.3 Influence of the task

We describe here other results testifying of the influence of the task performed on user preferences regarding factors influencing the SoE. For instance, we previously presented that the configuration the most

accepted was  $\{1,2,1\}$  in *Soccer*, *Fitness* and *Walking*, while  $\{2,2,1\}$  and  $\{3,2,1\}$  were the most accepted in *Punching*. It is interesting that in *Punching*, a task involving an interaction with the upper-body, reaching an equivalent SoE seems to require a higher level of appearance than for other tasks. A main difference between this task and the others is that participants have to look at their virtual body more closely in first-person PoV due to the proximity of the upper limbs with the PoV. While this could partially explain our result, we may also consider that interacting with a punching bag usually requires a consequent strength. It is therefore possible that stickman avatar did not fulfil the visual expectation of a strong body, compared to the realistic avatars, since it does not include any musculature. However, further research would be needed to better understand this result.

Another interesting finding is that in *Walking* the PoV was mostly increased first, rather than the control. Since the navigation task included obstacles on the ground, users were indirectly encouraged to focus on their feet. While it is difficult to say if it is the reason why participants tended to improve their PoV first, it is not the first time that differences in the way users perceived their avatar were highlighted for tasks involving a mirror or locomotion. Moreover, Koiliias et al. [20] showed that the way some factors influence the Sense of Agency varies depending whether the task performed is a self-observation task, an observation-through-mirror task or observation-during-locomotion task. It would therefore be interesting to explore different type of tasks, with or without mirror, and different types of navigation with or without obstacles, to further understand this result.

While it is hard to observe a specific pattern of influence depending of the task, results demonstrate that the preference between factors is not the same in all tasks. This outcome could question whether the results obtained in the baseline experiment were impacted by the specific task chosen. For this reason, it would be very interesting to further investigate whether the task influences the preference of levels of each factor independently, in addition to studying potential preferences in the way levels of different factors are combined together to achieve a satisfying SoE depending of the task.

## 7.4 Limitations and Future Work

The inter-relation between factors influencing the SoE is a complex process. While we tried to address in this paper the question of potential user preferences regarding these factors, we believe future research would be valuable to provide more insights on the subject.

In our paper, the choice of levels was constrained by the experimental design, where every level of a factor needed to be compatible with every level of all the other factors, as well as by technical limitations. In some cases, we may wonder how the limitations in implementation had an impact of user preferences. For instance, the limitation of our last level of appearance, i.e., the level of personalization of an existing 3D avatar, may explain partially why participants tended to accept configurations with low levels of appearance. Indeed, with such personalized avatars, the avatar body shape rarely matched the users', as well as the exact skin color. However, the current technological advancements are now starting to allow for the production of high-fidelity and highly-realistic avatars, as the ones created using photogrammetry [48] or seeking fidelity of body shape [35]. The addition of such a level of appearance, and more generally the use of a wider range of levels for each factor, would be valuable to more precisely evaluate user preferences and potential accepted configurations. For instance, while we decided in this first study to focus on only two levels of Point of View, inspired by the works of Gorisse et al. [16] and commonly used in video-games, more levels of Point of View could be considered, including for instance cinematographic aspects.

Furthermore, we may consider the potential influence of having different number of levels for each factor. Indeed, while we believe participants were aware that the same importance was to be given on each choice, their behaviour remains hard to control and we can not fully prevent the case of a user playing "optimally" by upgrading the factor that only takes one improvement. However, we did not observe stronger preferences for the point of view which had the lowest number of levels. We also believe that the choice of adding more levels should

be balanced by the fact that having too many levels without significant differences in terms of improvements could also lead to a different bias, hence the reason why a baseline experiment was conducted.

Moreover, while exploring appropriately more levels might still broaden the current findings, we also believe that including other factors in the process would be highly valuable. For instance, the present study only includes feedback about the visual aspect of the avatar. Future work could therefore consider exploring for instance the influence of a multisensory feedback factor, e.g., involving tactile and haptics feedback which are also known to influence the SoE [11, 22].

While the subjective matching method used in this paper enabled the exploration of factors influencing the SoE without the use of subjective questionnaires or behavioral measurements, we believe it is important to discuss the potential source of unreliability it may contain. First, the subjective matching technique enables the manipulation of a high number of levels and factors in one experiment. However, it also brings the risk of overloading the cognition of participants with all the configurations to remember. We may also consider the difficulty for participants to remember their SoE in the optimal configuration in order to match it from minimal avatar configurations, but also their potential difficulty in understanding the definition of SoE in the first place. Indeed, a description of the SoE to participants with less abstraction would ensure a better uniformity of what this feeling refers to among participants. However, the ratings given to each factor at the end of the experiments by participants, as well as some of their final comments regarding the experiment, testify of a certain guarantee in their choices during the subjective matching experiment: "Appearance is the less important aspect. I preferred the body with spheres or the second avatar", "The control for me was the most important factor, without control it really felt I was looking like someone else". Second, another limitation of this method is that when participants accepted lower configurations of avatars, we have no certainty that their SoE was indeed the same as the one felt in the optimal configuration. It would be interesting to assess participants SoE through subjective questionnaires right after participants accepted a lower configuration and after the optimal configuration. However, if this would be done for every trial of an experiment, the additional time added to the whole experiment would have to be addressed as a potential bias source.

## 8 CONCLUSION

In this paper, we presented two experiments exploring user preference and perception of three factors commonly found in the literature to influence the Sense of Embodiment in Virtual Reality, namely: the avatar's visual appearance, the avatar's control, and the user point of view. Our results first show that appearance of the avatar was given less importance than control or point of view. Second, we found that when it comes to virtual embodiment users do not necessarily need to reach the optimal avatar configuration to feel a fulfilling SoE, suggesting that VE designers may not always need to provide high-end graphics avatars but should provide a high degree of control. Third, we showed that the accepted configurations can vary depending on the task performed, stressing the importance of this aspect for future studies and applications. Taken together, our results provide valuable insights for designers of VR applications involving avatars, showing which factors among the three studied should be prioritized, and paving the way to future studies aiming at better understanding the inter-relations between factors influencing the Sense of Embodiment.

## ACKNOWLEDGMENTS

We wish to thank participants of our experiment. This work was sponsored by the Region Bretagne and the Inria Avatar Challenge.

## REFERENCES

- [1] F. Argelaguet, L. Hoyet, M. Trico, and A. Lecuyer. The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *2016 IEEE Virtual Reality (VR)*, pp. 3–10, March 2016. doi: 10.1109/VR.2016.7504682

- [2] J. E. Aspell, B. Lenggenhager, and O. Blanke. Keeping in touch with one's self: Multisensory mechanisms of self-consciousness. *PLOS ONE*, 4(8):1–10, 08 2009. doi: 10.1371/journal.pone.0006488
- [3] A. S. Azevedo, J. Jorge, and P. Campos. Combining eeg data with place and plausibility responses as an approach to measuring presence in outdoor virtual environments. *Presence*, 23(4):354–368, Nov 2014. doi: 10.1162/PRES.a.00205
- [4] I. Bergström, S. Azevedo, P. Papiotis, N. Saldanha, and M. Slater. The plausibility of a string quartet performance in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 23(4):1352–1359, April 2017. doi: 10.1109/TVCG.2017.2657138
- [5] O. Blanke and T. Metzinger. Full-body illusions and minimal phenomenal selfhood. *Trends in Cognitive Sciences*, 13(1):7–13, 2009. doi: 10.1016/j.tics.2008.10.003
- [6] P. Bourdin, I. Barberia, R. Oliva, and M. Slater. A virtual out-of-body experience reduces fear of death. *PLOS ONE*, 12(1):1–19, 01 2017. doi: 10.1371/journal.pone.0169343
- [7] E. A. Caspar, A. Cleeremans, and P. Haggard. The relationship between human agency and embodiment. *Consciousness and Cognition*, 33:226–236, 2015. doi: 10.1016/j.concog.2015.01.007
- [8] F. de Vignemont. Embodiment, ownership and disownership. *Consciousness and Cognition*, 20(1):82–93, 2011. Brain and Self: Bridging the Gap. doi: 10.1016/j.concog.2010.09.004
- [9] H. H. Ehrsson, K. Wiech, N. Weiskopf, R. J. Dolan, and R. E. Passingham. Threatening a rubber hand that you feel is yours elicits a cortical anxiety response. *Proceedings of the National Academy of Sciences*, 104(23):9828–9833, 2007. doi: 10.1073/pnas.0610011104
- [10] C. Farrer, M. Bouchereau, M. Jeannerod, and N. Franck. Effect of distorted visual feedback on the sense of agency. *Behav Neurol*, 19(1-2):53–57, 2008. 18413918[pmid]. doi: 10.1155/2008/425267
- [11] J. Frohner, G. Salvietti, P. Beckerle, and D. Prattichizzo. Can wearable haptic devices foster the embodiment of virtual limbs? *IEEE Transactions on Haptics*, pp. 1–1, 2018. doi: 10.1109/TOH.2018.2889497
- [12] H. Galvan Debarba, S. Bovet, R. Salomon, O. Blanke, B. Herbelin, and R. Boulc. Characterizing first and third person viewpoints and their alternation for embodied interaction in virtual reality. *PLOS ONE*, 12(12):1–19, 12 2017. doi: 10.1371/journal.pone.0190109
- [13] H. Galvan Debarba, S. Bovet, R. Salomon, O. Blanke, B. Herbelin, and R. Boulc. Characterizing first and third person viewpoints and their alternation for embodied interaction in virtual reality. *PLOS ONE*, 12(12):1–19, 12 2017. doi: 10.1371/journal.pone.0190109
- [14] M. Gonzalez-Franco and T. C. Peck. Avatar embodiment. towards a standardized questionnaire. *Frontiers in Robotics and AI*, 5:74, 2018.
- [15] M. González-Franco, D. Pérez-Marcos, B. Spanlang, and M. Slater. The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. In *2010 IEEE Virtual Reality Conference (VR)*, pp. 111–114, March 2010. doi: 10.1109/VR.2010.5444805
- [16] G. Gorisse, O. Christmann, E. Amato, and S. Richir. First- and third-person perspectives in immersive virtual environments: Presence and performance analysis of embodied users. *Frontiers in Robotics and AI*, 4:33, 07 2017. doi: 10.3389/frobt.2017.00033
- [17] G. Gorisse, O. Christmann, S. Houzangbe, and S. Richir. From robot to virtual doppelgänger: Impact of visual fidelity of avatars controlled in third-person perspective on embodiment and behavior in immersive virtual environments. *Frontiers in Robotics and AI*, 6:8, 2019. doi: 10.3389/frobt.2019.00008
- [18] P. Heidicker, E. Langbehn, and F. Steinicke. Influence of avatar appearance on presence in social vr. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 233–234, March 2017. doi: 10.1109/3DUI.2017.7893357
- [19] K. Kiltani, R. Groten, and M. Slater. The Sense of Embodiment in Virtual Reality. *Presence: Teleoperators and Virtual Environments*, 21(4):373–387, 2012. doi: 10.1162/PRES.a.00124
- [20] A. Kiliass, C. Mousas, and C.-N. Anagnostopoulos. The effects of motion artifacts on self-avatar agency. *Informatics*, 6:18, 04 2019. doi: 10.3390/informatics6020018
- [21] E. Kokkinara, K. Kiltani, K. Blom, and M. Slater. First person perspective of seated participants over a walking virtual body leads to illusory agency over the walking. *Scientific Reports*, 6:28879, 07 2016. doi: 10.1038/srep28879
- [22] E. Kokkinara and M. Slater. Measuring the effects through time of the influence of visuomotor and visuotactile synchronous stimulation on a virtual body ownership illusion. *Perception*, 43(1):43–58, 2014. PMID: 24689131. doi: 10.1068/p7545
- [23] B. Laha, J. N. Bailenson, A. S. Won, and J. O. Bailey. Evaluating control schemes for the third arm of an avatar. *Presence: Teleoperators and Virtual Environments*, 25(2):129–147, 2016.
- [24] M. E. Latoschik, D. Roth, D. Gall, J. Achenbach, T. Waltemate, and M. Botsch. The effect of avatar realism in immersive social virtual realities. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology, VRST '17*, pp. 39:1–39:10. ACM, New York, NY, USA, 2017. doi: 10.1145/3139131.3139156
- [25] B. Lenggenhager, T. Tadi, T. Metzinger, and O. Blanke. Video ergo sum: Manipulating bodily self-consciousness. *Science (New York, N.Y.)*, 317:1096–9, 09 2007. doi: 10.1126/science.1143439
- [26] L. Lin and S. Jörg. Need a hand?: How appearance affects the virtual hand illusion. In *Proceedings of the ACM Symposium on Applied Perception, SAP '16*, pp. 69–76. ACM, New York, NY, USA, 2016. doi: 10.1145/2931002.2931006
- [27] J. Lugin, M. Landeck, and M. E. Latoschik. Avatar embodiment realism and virtual fitness training. In *2015 IEEE Virtual Reality (VR)*, pp. 225–226, March 2015. doi: 10.1109/VR.2015.7223377
- [28] J.-L. Lugin. Anthropomorphism and illusion of virtual body ownership. 10 2015.
- [29] D. M. Wegner, B. Sparrow, and L. Winerman. Vicarious agency: Experiencing control over the movements of others. *Journal of personality and social psychology*, 86:838–48, 07 2004. doi: 10.1037/0022-3514.86.6.838
- [30] A. Maselli and M. Slater. The building blocks of the full body ownership illusion. *Frontiers in human neuroscience*, 7:83, 03 2013. doi: 10.3389/fnhum.2013.00083
- [31] A. Maselli and M. Slater. Sliding perspectives: dissociating ownership from self-location during full body illusions in virtual reality. *Frontiers in Human Neuroscience*, 8:693, 2014. doi: 10.3389/fnhum.2014.00693
- [32] S. Nagamine, Y. Hayashi, S. Yano, and T. Kondo. An immersive virtual reality system for investigating human bodily self-consciousness. In *2016 Fifth ICT International Student Project Conference (ICT-ISPC)*, pp. 97–100, May 2016. doi: 10.1109/ICT-ISPC.2016.7519245
- [33] M. Parger, J. H. Mueller, D. Schmalstieg, and M. Steinberger. Human upper-body inverse kinematics for increased embodiment in consumer-grade virtual reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology, VRST '18*, pp. 23:1–23:10. ACM, New York, NY, USA, 2018. doi: 10.1145/3281505.3281529
- [34] H. Preester and M. Tsakiris. Body-extension versus body-incorporation: Is there a need for a body-model? *Phenomenology and the Cognitive Sciences*, 8:307–319, 09 2009. doi: 10.1007/s11097-009-9121-y
- [35] S. Pujades, B. Mohler, A. Thaler, J. Tesch, N. Mahmood, N. Hesse, H. H. Bühlhoff, and M. J. Black. The virtual caliper: Rapid creation of metrically accurate avatars from 3D measurements. *IEEE Transactions on Visualization and Computer Graphics*, 25:1887, 1897, 2019.
- [36] D. Roth, J. Lugin, J. Büser, G. Bente, A. Fuhrmann, and M. E. Latoschik. A simplified inverse kinematic approach for embodied vr applications. In *2016 IEEE Virtual Reality (VR)*, pp. 275–276, March 2016. doi: 10.1109/VR.2016.7504760
- [37] M. V. Sanchez-Vives, B. Spanlang, A. Frisoli, M. Bergamasco, and M. Slater. Virtual hand illusion induced by visuomotor correlations. *PLOS ONE*, 5(4):1–6, 04 2010. doi: 10.1371/journal.pone.0010381
- [38] M. J. Schuemie, P. van der Straaten, M. Krijn, and C. A. van der Mast. Research on presence in virtual reality: A survey. *CyberPsychology & Behavior*, 4(2):183–201, 2001. PMID: 11710246. doi: 10.1089/109493101300117884
- [39] V. Schwind, P. Knierim, C. Tasci, P. Franczak, N. Haas, and N. Henze. "these are not my hands!": Effect of gender on the perception of avatar hands in virtual reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17*, pp. 1577–1582. ACM, New York, NY, USA, 2017. doi: 10.1145/3025453.3025602
- [40] R. Skarbez, S. Neyret, F. P. Brooks, M. Slater, and M. C. Whitton. A psychophysical experiment regarding components of the plausibility illusion. *IEEE Transactions on Visualization and Computer Graphics*, 23(4):1369–1378, April 2017. doi: 10.1109/TVCG.2017.2657158
- [41] M. Slater, B. Spanlang, and D. Corominas. Simulating virtual environments within virtual environments as the basis for a psychophysics of presence. *ACM Trans. Graph.*, 29(4):92:1–92:9, 2010. doi: 10.1145/1833351.1778829
- [42] M. Slater, B. Spanlang, M. V. Sanchez-Vives, and O. Blanke. First person experience of body transfer in virtual reality. *PLOS ONE*, 5(5):1–9, 05

2010. doi: 10.1371/journal.pone.0010564
- [43] W. Steptoe, A. Steed, and M. Slater. Human tails: Ownership and control of extended humanoid avatars. *IEEE Transactions on Visualization and Computer Graphics*, 19(4):583–590, April 2013. doi: 10.1109/TVCG.2013.32
- [44] M. Synofzik, G. Vosgerau, and A. Newen. Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*, 17(1):219 – 239, 2008. doi: 10.1016/j.concog.2007.03.010
- [45] M. Tsakiris, G. Prabhu, and P. Haggard. Having a body versus moving your body: How agency structures body-ownership. *Consciousness and Cognition*, 15(2):423 – 432, 2006. doi: 10.1016/j.concog.2005.09.004
- [46] E. Tuvér, L. Macis, F. Sorrentino, L. D. Spano, and R. Scateni. Fitmersive games: Fitness gamification through immersive vr. In *Proceedings of the International Working Conference on Advanced Visual Interfaces, AVI ’16*, pp. 212–215. ACM, New York, NY, USA, 2016. doi: 10.1145/2909132.2909287
- [47] M. Usoh, E. Catena, S. Arman, and M. Slater. Using presence questionnaires in reality. *Presence: Teleoperators and Virtual Environments*, 9, 04 2000. doi: 10.1162/105474600566989
- [48] T. Waltemate, D. Gall, D. Roth, M. Botsch, and M. E. Latoschik. The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1643–1652, 2018. doi: 10.1109/TVCG.2018.2794629
- [49] T. Waltemate, I. Senna, F. Hülsmann, M. Rohde, S. Kopp, M. Ernst, and M. Botsch. The impact of latency on perceptual judgments and motor performance in closed-loop interaction in virtual reality. In *Proceedings of the 22Nd ACM Conference on Virtual Reality Software and Technology, VRST ’16*, pp. 27–35. ACM, New York, NY, USA, 2016. doi: 10.1145/2993369.2993381
- [50] L. Wood, H. Teräs, T. Reinert, and S. Gregory. The role of gamification and game-based learning in authentic assessment within virtual environments. In *Proceedings of HERDSA Conference*, 07 2013. doi: 10.13140/2.1.3890.5282